

A Lifetime of Language Experience at Each Click!

PAC3 at JALT 2001

Conference Proceedings



International Conference Centre

Kitakyushu JAPAN

November
22-25, 2001

Ramesh Krishnamurthy
COBUILD (Collins Dictionaries
and University of Birmingham)

Native speakers learn their mother tongue slowly, from birth, by the constant repetition of common words and phrases in a variety of contexts and situations, within the language community. As foreign language learners, we face considerable disadvantages when compared to children learning their mother tongue. Foreign language learners start later in life, have less time, have fewer opportunities to experience the language, and learn in the restricted environment of the classroom. Teachers and books give us information about many words and phrases, but it is difficult for us to know what we need to focus on and learn thoroughly, and what is less important. The rules and explanations are often difficult for us to understand.

A large language corpus represents roughly the amount and variety of language that a native-speaker experiences in a whole lifetime. Learners can now access language corpora. We can check which words and phrases are important, and quickly discover their common meanings, collocations, and structural patterns. It is easier to remember things that we find out ourselves, rather than things that teachers or books tell us. Each click on the computer keyboard can show us the same information in different ways, so we can understand it more easily. We can also get many more examples from a corpus. Teachers and native-speakers can also use corpora, to confirm and enhance their own knowledge of a language, and prepare exercises to guide their students. Each of us can learn at our own level and at our own speed.

The need for English

In the June, 2001 issue of *The Language Teacher*, David Nunan said, “The demand for English language education has exploded with economic globalization”. He mentioned several fields in which English is now dominant: business, technology, science, the Internet, popular entertainment, and sports. He quoted Swales’ 1987 estimate that over 50% of academic papers published were in English, and the percentage has grown every year since then. And Nunan also mentioned that several countries are now introducing the study of English at primary school level.

Economic globalization continues to expand. China has just become a member of the World Trade Organization, adding another 1.3 billion people to the global market. I was in Shanghai recently, and the new National College English Course will be compulsory for all students who are not English majors. This means that over 20 million students will be taking the course. So, unquestionably, the demand for English is still increasing worldwide.

The differences between learning L1 and learning L2

L1 learners benefit from constant language repetition right from the start, from their family and from the wider community. They also have strong motivation: language is part of the bonding with parents, the gradual

emergence of a sense of individual identity, as well as the means to joining in social activities. Most important of all, L1 learners are given constant praise for every word or phrase they master.

Language reinforcement is also provided by the L1 environment, all day and every day, while playing with friends, watching television, or going shopping. Street signs, advertisements, and notices are seen, and snippets of conversation are overheard. Whether the process takes place at a conscious or a subconscious level, language reinforcement is taking place. L1 learners come across the language in a wide variety of contexts, and constantly “triangulate” words and phrases across the different discourses, language modes and registers. The L1 learning process takes place over a long timescale, so there is less anxiety. Both teachers and learners are relaxed for the most part.

L2 learners, on the other hand, suffer in every single one of these areas. There are usually gaps in the repetition of words and phrases: we may see or hear a new word today and not come across it again for a week, a month, or a year. By then, we are learning it as a new word again, having forgotten all about the previous encounter. Motivation is also much weaker: rather than being personally and communally cohesive, learning an L2 may even be socially alienating, there is certainly less praise and more error correction.

Reinforcement is likewise minimal and sporadic,

perhaps only occurring in the language classroom once or twice a week. L2 learners experience a greatly reduced number of varieties, genres, and registers: the classroom context is often the only source. During self-study, the environment is even more isolated. The L2 learning process is, in general, accorded a shorter timescale, and the setting is nearly always more formal. Both teachers and books are trying to cram in more information into the abbreviated schedule, so there is less opportunity to decide what is important, what is relevant, etc. Less guidance is available, from fewer sources. Hence the activity is prone to anxiety, stress, and tension.

L1 learning is a more inductive process. Learners are given substantial amounts of appropriate input, and gradually infer a pattern or rule from the evidence. Learning L2 involves more deductive processes. We are often given the rules and generalizations first, with an example or two afterwards. Abstract rules are difficult to learn, because they do not have as much psychological impact. Usually, we get an insufficient number of examples to get us to the critical mass of evidence that is required to make the rule comprehensible. Doubts and queries arise frequently and the majority of them remain unresolved.

Dictionaries and other reference resources

For most L2 learners, the dictionary is the first and most regularly used reference resource. L1 learners rarely use

reference works until much later in the learning process. Dictionaries, grammar books, and course books can be very useful for some purposes and on some occasions. However, dictionaries and grammar books aim at a level of generalization far removed from the specific complexities of context. Course books tend to make points that are limited in applicability to the text or passage in question. Neither provides adequate help to the learners who are struggling with the particular text they are studying or trying to create.

Let us look at what actually happens when we consult a reference work. For example, we may be reading a text, and trying to guess the meaning of a word from its current context, trying to recall vague memories of previous encounters with this word, etc. This is an inductive process. Then we decide to look up the word in our dictionary. We suddenly have to switch to a deductive process, working with higher-level information about the word, which is embedded in a template of lexicographic conventions.

Similarly, if we are trying to write an essay, we are creating a context. If we stop to consult a dictionary or grammar book, we are diverted from our intended context and have to try to match our intended message with a world of secondary information (*about* the word) that is very differently encoded.

Language corpora

Now imagine the difference if, when we are reading or writing, and come across a problem or doubt, we could immediately see several examples of the word or phrase we are stuck on. For example, I am reading a newspaper and find, below a picture of someone wagging a finger at a cat, the headline: “Enquiry warns fat cats to curb sminting excesses”. What does “sminting” mean?

A quick check in a publicly available corpus (I used the free 56 million word Collins Wordbanks *Online* English corpus at <http://www.cobuild.collins.co.uk/>) gives me the message “There were no instances of your search pattern found in the corpora you selected”. So I know that “sminting” is not an important word for me to learn. In fact, if I had looked at the newspaper article more closely, I would have noticed another picture, of a packet of mints called SMINT, and realized that the whole article was a humorous advertisement.

Let us take another example. We are writing and want to use *the authority* as the subject of the sentence. We know it is similar to collective nouns like *group* and *team*. But should we use a singular or plural verb? Should it be “The authority is ...” or “The authority are ...”? To find this information in a dictionary or grammar book would take some time, and the information might not be there, or might not be easily extracted. The same corpus as above shows us the

following examples:

WHO RUNS ITS? [/h] Ultimately the Authority is controlled by the Council who

this goal has been achieved and the Authority is pleased to announce that

the lowest levels recorded. [p] The Authority is committed to the provision of

for the West Midlands area. The Authority is funded partly through

Radio. In addition, the authority is advertising two new licences

supplies pose a health risk [p] The authority is right to hold an investigation

confusion of priorities. [p] The authority is no stranger to controversy.

[p] But at the same time the authority is splashing out £ 50,000

And when I type in the query “the+authority+are”, I again receive the message: “There were no instances of your search pattern found in the corpora you selected”. Problem solved! I must use the singular verb. But now I have another doubt. I want to talk about the authority which is responsible for health services in my town. Can I say “The authority for health...”? I noticed that the corpus website offers another facility: collocations. So I ask for the collocations for “authority”, and immediately

get a list of the top 100 collocates. The word “for” does not occur at all. Here are the top 10 collocates.

Collocate	Corpus Frequency	Joint Frequency	Significance
local	15400	619	24.515681
the	2872094	2942	23.098783
health	12755	431	20.399210
national	23881	161	11.581697
education	8325	128	10.880955
police	20057	104	9.041366
civil	3859	85	8.973379
aviation	441	81	8.971182
by	230325	276	8.459680
rivers	1013	72	8.415071

So, just to check, I go back to the examples facility (called the *concordance sampler*) and type in “authority+for”. To my confusion, I find many examples of this (the free access limits me to 40 examples each time):

na Mona, the Irish Peat Development Authority for accurate information

the said Land the District Planning Authority for the purposes of the Town and

by LAUTRO (the regulatory authority for the marketing of life

You need to contact your local authority for the latest details of the

Resolution 6 simply renews this authority for a further five-year period

by the Personal Investment Authority) for life assurance, pension and

by the Personal Investment Authority for investment business. [p]

kitchen table. Contact your local authority for courses in computer studies,

pound; 250 by the National Rivers Authority for not having a rod licence when

legal aid to sue her local health authority for prescribing the drug AZT to

However, I see that none of the words after “for” describes the function of the authority. And indeed I see that the last example, “local health authority”, is exactly what I need.

But just to make absolutely sure, I ask for examples for “local+health+authority” and find 40 of them (only 10 are shown here):

decision made by the local health authority and Virginia Bottomley, who was

and property from the local health authority by deception.

[p] The case

treatment, the local health authority can seek an order from a local

is under threat. The local health authority, Ealing, Hammersmith and

legal aid to sue her local health authority for prescribing the drug AZT to

London site, as the local health authority has guaranteed to cover the cost

Health Commission, the local health authority, has produced a catalogue of

meeting called by the local health authority last night, are threatening to

the reaction of the local health authority later. [p] The baleful effect of

Some lawyers said the local health authority might have the power to detain

This also confirms my decision to use a singular verb: two examples contain “the authority has” (rather than “the authority have”). One last check: “the+authority+have” gives the message: “There were no instances of your search pattern found in the corpora you selected”, whereas “the+authority+has” gives me:

family as it is seen today. [p] The authority has a clear mandate to design and

However, the person delegating the authority has still to

keep overall

excessive effluent. [p] The authority has since applied to have

the issue of unlawful payments, the authority has a duty to ensure public money

What do you believe [p] But the authority has done its best. It has at

There are other free access corpora (e.g. the British National Corpus Online facility “simple search” at <http://sara.natcorp.ox.ac.uk/lookup.html>). But full access (usually obtained by paying a subscription, or buying a copy of the corpus) allows much more interesting and detailed searches, once we get used to the software system. We can get frequency lists, genre information, grammatical analyses, and so on, so students and teachers can study a wide range of linguistic features in a large variety of texts.

Another way to begin using corpora is to create your own, which is fairly easy with so much data available on the Internet. You can then buy some good, easy-to-use corpus software (such as Wordsmith Tools or Monoconc) and do quite complex analyses with smaller amounts of data. This is particularly useful for students and teachers working in ESP or EAP.

Corpora can also be easy sources of exercises for the classroom. When students come across an unknown

word in a reading passage, we often tell them to try to guess the meaning of the word from its context. Yet we do not give them much practice or training in using context. One easy exercise to prepare is “gapped concordances”. The keyword is deleted from a set of examples, and students not only have to guess the missing word, but give a step-by-step account of what clues in the contexts made them decide on that word.

to have a run at it. He's very and has loads of abi
from settled. <p> Sinclair, 22, is , brave, composed and
EC negotiators warned that if no changes to reinforce
left outdoors untreated. However, a coating of teak oil
for a free election, including a ` count” by citizen mo
campers and students who want a cuppa but don't
have
as a service to clients who needed delivery of securiti
the car he leaned over and gave me a peck, a goodnight
ki
to conclude may I just ask you a few questions about
your
the bombs were falling.” Trish was to support Terry's
s Lord of the Rings. Others have been to follow but the
BB
so easy. It seems to be so simple and to take a big
brush

What part of speech is the missing word? It must be an adjective. Why? Because it comes after the verb *be*, comes after the adverb *very*, and comes between the indefinite article and a noun. What other features can we notice? The missing adjective describes people, and also things (*changes, count, cuppa, delivery, peck, questions*; students could be asked to look up the collocates of these words, to see if they share the same collocate, which might be a candidate for the missing word). The adjective is followed by a to-infinitive in the last 3 examples (so students could be asked to look in a grammar book to find a list of adjectives which have this feature, or the teacher can give them a list of such adjectives). In some cases, longer contexts might help, and this can be made available by most corpus software. In this case, the missing word is *quick*.

Teachers can introduce new words by showing a small selection of examples, and highlighting the common collocations and grammatical patterns:

has no previous record as **an authority on** Islamic affairs

As **an authority on** medieval church archit

A leading authority on the geology of Stonehe

Bears **stamped their authority on** the game in the first

his intention to **stamp his authority on** the Scottish Tories

marine officer used to exerting **authority over** men, thought that

Liberal leader had the kind of **authority over** the party that Howar

partner holds a position of **authority over** the younger person

that if I could have had that **authority over** them I probably woul

but only Congress has the legal **authority to declare** war

But the force will have no **authority to open fire**

the CJC had no lawful **authority to investigate**

Variations in spelling, in genre, and in meaning can become much clearer if we see a few examples of each. For example, we can look at some examples for *womanly* and *womanish*:

when women were supposed to look womanly. [p] Just a generation ago, the

be getting well, you know, even more womanly. At least then the male sex could

luxury is traditionally about a womanly boudoir filled with silk, satin

antifeminism, they presumed that a womanly conception of politics was the

The 39-year-old Hollywood star whose womanly curves once wooed a generation of

performance of unspecified future womanly duties. 1837 Aug. 5 [p] The first

fluidly over her body, revealing the womanly fullness over her hips for the

I accept my body with all its womanly fullness. [p] Let the mirror echo

Her boyish awkwardness turns to womanly grace. And she's in on in the

thin, she added, the figures had womanly hips and breasts. [p] Emily

for policy; the anxiety over womanish traits--one's own or those of one'

as a fairy - a 'lisper, swishing, womanish-acting man," in one of Bessie

morally debased.⁷⁹ Similarly, the 'womanish-acting man' became a special

pejorative: 'feminine wiles # womanish tears # a womanlike lack of

dictionary, after defining the word womanish as 'suitable to or resembling a

the negative images associated with womanish and womanlike, and with

Those old-timers were rather womanish, some of them, he thought. All a

manly or less than acceptable as womanish. I mean what's wrong with being a

skin, accentuate the almost womanish bow of his mouth.

The incurable

He was the sort of slightly womanish man who tended to be drawn towards

the dream so popular with women and womanish men of a magnanimous bully. <0> It

Even from these few lines, it will be seen that *womanly* is a positive way of describing a woman, and *womanish* is a negative way of describing a man. More examples and a closer study shows that *womanly* tends to describe a woman's body, whereas *womanish* describes a man's behaviour. This can be the starting point for a study of words describing physical appearance and words describing behaviour, or words used to describe women as opposed to words used to describe men.

Conclusions

In this way, teachers can help L2 learners to become used to seeing words in a variety of texts and contexts, to make them recognize collocations and grammatical patterns, and to understand how contexts can contribute to differences of meaning. The wealth of typical examples can partly make up for the lack of L2 language repetition and reinforcement in an L1 culture. The use of computers and new forms of exercises can boost the students' motivation. Learners can use the corpora as much or as little as they wish. They can also gain access to corpora outside normal lessons, and indeed at any time. Some queries, studies, or exercises can be completed within a few minutes, others can take longer. Each click on the keyboard offers another glimpse into the language and how it works. Some patterns are easy to spot, others need more careful attention. With appropriate preparation, practice, and guidance, corpora can allow each of us to learn in our own way, at our own level, and at our own speed.